From Signature-Based Towards Behaviour-Based Anomaly Detection (Extended Abstract)

Pavel Minarik, Jan Vykopal
Masaryk University
CZECH REPUBLIC

minarik@ics.muni.cz / vykopal@ics.muni.cz

INTRODUCTION

It has been an continuous phenomenon that more and more information is transmitted and accessible via computer data networks. Therefore data networks become a critical spot with lots of risks and threats related to it. One example can be a temporary dysfunction of network caused by an intended attack (such as DDoS attack). Attacks may lead to server failures which can mean simple inability to provide required services but also they can paralyse systems on national level (what recently happened in Estonia and Georgia [1]). Another example of possible thread is a loss of credibility of data, e.g. by unauthorized access and manipulation with stolen data. Crucial elements of data network can be overpowered by an attacker, for instance by breaking down password and setting administration access rights. Result of such activity can end up by misusing the element of data network for illegal actions (e.g. phishing, botnet) or by continuous abuse of the network.

STATE OF THE ART

Generally, we can divide intrusion detection systems (IDS) into two basic classes according to their position in the network: host-based intrusion detection systems and network-based intrusion detection systems. We believe that network-based IDS (NIDS) has the following advantages: In contrast to Host-based IDS (HIDS), the deployment of a new host in network does not demand more effort to monitor the network activity of the new host. There is no need to install any specialized software on the host. Network may consist of some specialized hosts (besides common servers or workstations).

So, the HIDS installation is impossible in such a case. Next, NIDSs are passive devices, “invisible” for the attackers. On the contrary, HIDSs rely on processes that running in the operating system of the host. We also consider the deployment, testing and possible upgrade of IDS. Generally, it is easier to update one component of NIDS than many components of HIDS on hosts.

Many systems used for a defence of cyber threats are based on the most common approach so called Deep Packet Inspection or L7 Decoding. Deep Packet Inspection approach consists in analysis of packet contents and brings good results for general and therefore less serious attacks. For professionally prepared attacks coming from inside of the network results of this method are significantly less powerful. There are also some other constraints when dealing with the content of data packets. The methods of payload analysis are very demanding for network performance and cannot be used in encrypted traffic while the ratio of encrypted traffic is increasing. An alternative approach is a Behaviour Analysis which uses information from the L3/L4 layer (i.e. characteristics of data flows in IP networks) and does not work with the content of packets at all. A combination of both methods ensures a higher ability of system to react on a wider scope of threads and therefore increases security of a network in general as we will demonstrate in this paper.

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **NOV 2010**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **From Signature-Based Towards Behaviour-Based Anomaly Detection (Extended Abstract)**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Masaryk University CZECH REPUBLIC**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

See also ADA564697. Information Assurance and Cyber Defence (Assurance de l'information et cyberdefense). RTO-MP-IST-091

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**SAR**	**4**	

DEEP PACKET INSPECTION

Every packet and even its payload needs to be inspected. The core of the signature-based detection is generally “expensive” string matching. This is the essential limitation in high-speed (multigigabit) networks. For example, it is impossible to run well-know network IDS Snort on COST (commercial off-the shelf) hardware without any packet loss even on 1 Gigabit Ethernet. This limitation can be overcome by specialized hardware cards (produced e. g. by Endace, Napatech or Invea-tech) that accelerate packet capture. However, these cards are quite expensive. Last, but not least, deep packet inspection often works only with “local” information (only with the packet that is currently passing the device). Finally, there is another important limitation of Deep Packet Inspection. This method is completely useless in case of encrypted traffic (payload encryption, IPSec tunnels, etc.).

NETWORK BEHAVIOUR ANALYSIS

The classic approach of many IDS or IPS to data collection is to capture all network packets that pass through the system. In contrast, network behaviour analysis relies on information and statistics of network flows.

Many routers and monitoring probes that perform flow-based data acquisition can serve as sensors. De-facto standard for IP flow monitoring is NetFlow format. Although NetFlow was originally developed by Cisco Systems (version 5), the latter was standardized as an open protocol (version 9) by IETF in 2006. A flow is defined as an unidirectional sequence of packets with some common properties that pass through a network device. These collected flows are exported to an external device, the NetFlow collector. Network flows are highly granular; for example, flow records include details such as IP addresses, packet and byte counts, timestamps, Type of Service (ToS), application ports, input and output interfaces, etc. [2] Thus, the flow acquisition provides an aggregated view of network traffic and typically do not provide any information about payload.

What is more, it significantly reduce amount of data that need to be processed by methods of network behaviour analysis.

There are several types of currently used methods of Network Behaviour Analysis:

- Statistical methods that find anomalies in terms of clusters and outliers or time series (e. g. [3])
- Continuous host profiling with interest in changes of behaviour.
- Heuristics are focused at more or less general traffic patterns.

BEHAVIOUR ANALYSIS APPLICATIONS

There is a plenty of network security related issues. None of the two main presented approaches solves all of them. In previous text we paid attention to the main differences between Deep Packet Inspection and Behaviour Analysis. In this section we show examples of utilization of these methods for various task related to network security.

Task	Deep Packet Inspection	Behaviour Analysis
Application protocol analysis	YES	NO
Signature-based IDS	YES	NO
Peer-to-peer networks	YES	YES
Dictionary attacks detection	NO	YES
Host profiling	NO	YES
Unknown threats	NO	YES

Application protocol analysis is a common task of Deep Packet Inspection where a packet payload is checked for specific patterns to detect particular network application or service. Some significant characteristics of related flows may be checked by Behaviour Analysis but in general it is not an applicable method for this task. Another example of Deep Packet Inspection utilization is a **signature-based intrusion detection** where each packet is checked for a specific signature (pattern) occurrence.

Peer-to-peer networks are typical example of undesired traffic in corporate networks. They are widely used to share illegal content and for illegal data transfers. Both Deep Packet Inspection and Behaviour Analysis may be applied here. While Deep Packet Inspection looks for a particular pattern (application of a signature-based IDS), Behaviour Analysis evaluates flow characteristics and looks for typical behaviour of Peer-to-peer client communication (small and short connections to locate peers followed by massive data transfers).

Dictionary attacks are the most “favourite” and widely spread forms of attacks. The aim of a dictionary attack is to obtain an unauthorized access to a service, data or even to a network device. A typical example of dictionary attack is an attack led against a SSH server. Based on our experience with the university network defence we developed a detection algorithm (adaptive heuristic) based on a generic SSH authentication pattern [4]. Thanks to the network-based approach using NetFlow data, the detection algorithm is host independent and highly scalable. Deep Packet Inspection approach cannot detect dictionary attacks while the SSH traffic is encrypted.

A host profiling and detection of changes of behaviour (an anomaly detection) is a new task directly developed from flow monitoring and analysis. Deep Packet Inspection based methods are not suitable for this task at all because significant characteristics of flows are the amount of traffic, used services, provided services or communication peers not packet contents. We contributed to this topic by [5].

Behaviour Analysis is the key to react even to **unknown threats**. Deep Packet Inspection needs to know the attack signature but how to provide signature of an unknown threat? However, Behaviour Analysis, especially statistical methods or host profiling, may indicate a significant change of host behaviour and signalize undesired situation (e.g. zero day attack).

CONCLUSION

In this extended abstract, we presented a brief overview of two main approaches to deal with network security issues. We showed some examples where behaviour-based analysis has a large potential and can complement a traditional signature-based approach. It is capable to deal with issues where Deep Packet Inspection is inappropriate. Nevertheless, there is no almighty approach and the key is a combination of both signature-based detection and behaviour-based detection methods.

REFERENCES

[1] "Georgia targeted in cyber attack", AFP news, Retreived online: http://afp.google.com/article/ALeqM5iRuGsssizXAKVgmPqAXOxqB5uHsQ.

[2] Claise, B. (Ed.): RFC 3954: "Cisco Systems NetFlow Services Export Version 9", 2004, Retreived online: http://www.ietf.org/rfc/rfc3954.txt.

[3] Ertöz, L. and Eilertson, E. and Lazarevic, A. and Tan, P. and Kumar, V. and Srivastava, J. and Dokas, P.: "The MINDS - Minnesota Intrusion Detection System". In Next Generation Data Mining, Boston: MIT Press, 2004.

[4] Vykopal, J. and Plesnik, T. and Minarik, Pavel.: "Network-based Dictionary Attack Detection". In Proceedings of International Conference on Future Networks (ICFN 2009, Bangkok). Los Alamitos, CA, USA : IEEE Computer Society, 2009. pp. 23-27. ISBN 978-0-7695-3567-8.

[5] Minarik, P. and Krmicek, V. and Vykopal, Jan.: "Improving Host Profiling With Bidirectional Flows". In 2009 International Conference on Computational Science and Engineering. Vancouver, Canada : IEEE Computer Society, 2009. pp. 231-237. ISBN 978-0-7695-3823-5.